# 童年中期儿童执行功能的统一性和多样性：基于潜在变量分析和网络分析[1]

黄碧娟[1, 2] 朱红祥[1] 刘畅[1] 雒瑞帆[1] 封洪敏[1] 李红霞[1] 司继伟[1, *]
1 心理学院，山东师范大学，济南，250358
2 脑科学与类脑研究院，山东第一医科大学（山东省医学科学院），济南，2550117

**摘 要** 作为人类的高级认知功能，执行功能是促进个体学业成功和身心健康发展的核心要素。然而当前相关研究对执行功能的衡量侧重单一任务，缺乏从整合视角考察其结构及稳定性。本文采用纵向设计，对 756 名三至四年级儿童（平均年龄 9.25，女生 51.85%）进行一年追踪调查，通过潜在变量分析和网络分析，探究童年中期儿童执行功能的结构及稳定性。结果显示：童年中期儿童的执行功能以抑制控制和认知灵活性/工作记忆共同因子（即“I + W/C”模型）的形式存在；6 项执行功能任务被识别为一个稳定的成分；并且该期间执行功能的统一性（一般因子）和多样性（成分相关因子）相对稳定。上述发现强调了同时关注执行功能统一性和多样性的必要性，推进对儿童执行功能结构发展的深刻认识。

**关键词** 执行功能，统一性，多样性，童年中期
**分类号** B844

# Unity and Diversity of Children’s Executive Functions During Middle Childhood: Latent-Variable Analysis and Network Analysis

Bijuan Huang[1, 2] Hongxiang Zhu[1] Chang Liu[1] Ruifan Luo[1] Hongmin Feng[1] Hongxia Li[1] Jiwei Si[1, *]
1 School of Psychology, Shandong Normal University, Jinan, 250358, China
2 Institute of Brain Science and Brain-inspired Research, Shandong First Medical University & Shandong Academy of Medical Science, Jinan, 2550117, China

**Abstract** As high-level cognitive processes, executive functions (EF) refer to a set of top-down neurocognitive processes served for conscious, goal-directed control of thought, action and emotion, which is crucial for children’s academic success and mental and physical health. It includes inhibitory control, working memory and cognitive flexibility. Although abundant related theories and studies existed, some limitations still remained. Firstly, EF are measured with a single task in previous studies, which are rarely incapable of decomposing different components of EF. They are lack to systematically investigate the unity and diversity of EF from an integrative view; Then, rare studies use longitudinal design to examine the stability of the unity and diversity of children’s executive functions during middle childhood. The present attempted to address these issues mentioned above.

From the perspective of a unity/diversity framework of EF, combined with latent-variable analysis and network analysis, the present study followed 756 students from grade 3 to 4 ($M_{age}$ = 9.25 years, 51.85% girls) to systematically and comprehensively explore the unity and diversity of EF during middle childhood. They were recruited to measure six tasks related with three main components of EF.

The results showed that: The structure of children’s EF in middle childhood included inhibitory control and combined factor of working memory and cognitive reflection; These six

[1] 本文系国家自然科学基金面上项目“儿童数学焦虑变化的个体-环境交互作用：基因-脑-行为研究”（项目编号：31971010）的研究成果之一。

tasks were organized into one stable components through at least age 10; The unity and diversity of EF met full metric invariance.

These findings highlight the importance of understanding the unity and diversity of EF among middle childhood from a developmental perspective, and provide new sight on the measuring means of EF.

**Key words** executive functions, unity, diversity, middle childhood

## 1 引言

执行功能（Executive Functions）也称执行控制或认知控制。相较于认知控制，执行功能这一概念出现更早，而且与传统的神经心理学框架更为契合。执行功能指完成对思想、行为和情绪有意识的、目标导向控制时的一系列自上而下的神经认知加工过程（Zelazo & Carlson, 2012）。它是人类核心的高级功能之一，是个体学业成功和身心健康发展的核心要素（Huizinga et al., 2006）。执行功能涵括许多心理过程，这些过程允许个人利用思想控制行为，并进行复杂的活动，例如计划、组织、控制、维持注意力和自我管理（Goldstein & Naglieri, 2014）。一般认为，执行功能存在三个核心成分或功能，即抑制控制（Inhibit Control）、工作记忆（Working Memory）和认知灵活性（Cognitive Reflection）（Miyake et al., 2000）。

### 1.1 执行功能的测量

目前有两类常见的执行功能评估方式：基于表现的测试（如，抑制控制通过Go/Nogo 任务评估）和标准化问卷测评（如，BRIEF）（Isquith et al., 2013; McAuley et al., 2010; Strauss et al., 2006; Toplak et al., 2013）。接下来，本文将对以上两类主要评估方式进行简要介绍。

标准化问卷测评工具主要包括 BRIEF、BDEFS-CA、D-REF、BASC-2 和 CEFI 等（Goldstein & Naglieri, 2014）。其中，BRIEF 是目前最常用的标准化问卷测评。BRIEF-2 适用于评估 5 ~ 18 岁儿童的执行功能行为表现，由父母或教师根据儿童的日常行为表现进行评估，测量儿童的行为调节（抑制、转换和情绪控制）和元认知问题解决（启动、任务组织/计划、环境组织、自我监控和工作记忆）。然而，标准化问卷测验并不是评估儿童执行功能的最佳测验形式，原因具体如下：一方面，教师或父母报告的执行功能并不能有效反应儿童的真实状况；另一方面，标准化问卷测评题项数量相对较多，儿童的语言能力有限，并且回答动机较差（Shi et al., 2019）。儿童的执行功能测验需简单易懂，相对独立于语言能力，足够简化以保持儿童动机，甚至有趣。

与标准化问卷测评相比，基于表现的测试更适用于评估儿童的执行功能。基于表现的测试关注一般能力，更精确地记录儿童反应的速度和正确性，降低对语言指导和反馈的依赖（De Luca et al., 2003; Wild & Musser, 2014）；此外，它还可捕捉执行功能的特定成分。当探讨执行功能发展的理论问题时，基于表现的测试是必要的（Shi et al., 2019）。基于表现的测试包含行为测验集（Battery）和单个测验任务（Individual Task）（McCoy, 2019）。研究中常用的测验集包括 CANTAB、D-KEFS、TEC 和 NIHTB-CB 等（Goldstein & Naglieri, 2014）。其中，NIHTB-CB是一套比较全面的神经行为测量方法，可快速评估认知、情感、感觉和运动功能，适用于 3 ~ 85 岁被试。单个测验任务旨在关注执行功能的某个成分，工作记忆的测验任务主要有韦氏智力测验中的倒序数字广度（Backward Digital Span）、N-back 范式等；抑制控制的测验任务主要有 Oddball、Go/Nogo、Stop-Signal 等范式；认知灵活性的测验任务主要有 WCST（Wisconsin Card Sorting Test）、TMT（Trial Making Test）等。

上述两类评估方式提供了不同但重要的信息（Toplak et al., 2013）。鉴于本文考察的对象为童年中期儿童，以及研究重点是考察儿童执行功能的结构及稳定性，最终选用基于表现的测试评估儿童的执行功能。

### 1.2 执行功能结构的构建

在一些采用多任务衡量执行功能的研究中，虽然任务之间存在较低或不显著的相关（Huizinga et al., 2006），但潜在变量分析可以用来识别所观察到的认知任务表现的潜在结构（Gorsuch, 1983; Karr et al., 2018; Wiebe et al., 2008）。例如，探索性因子分析可析出多个潜在的因子结构（Friedman & Miyake, 2017; Karr et al., 2018; Miyake et al., 2000）。在评估个体执行功能时，基于表现的测试会面临一个主要挑战——任务杂质（Task Impurity），即大多数测试涉及非执行功能加工过程（Karr et al., 2018; Miyake & Friedman, 2012; Miyake et al., 2000; Toplak et al., 2013）。例如，Stroop 任务中的颜色命名和发音速度（Miyake & Friedman, 2012）。采用验证性因子分析（Confirmatory Factor Analysis, CFA）抽取这些任务的共同方差，可将任务杂质最小化（e.g.,Huizinga et al., 2006; Miyake et al., 2000）。通过提取较为“干净”的潜在变量，验证性因子分析构建和识别出来的成分被认为是不同任务表现的基础认知结构（Miyake et al., 2000）。越来越多的发展研究采用验证性因子分析探讨学前儿童（e.g., Carlson et al., 2014; Clark et al., 2013; Espy et al., 1999; Garon et al., 2008; Wiebe et al., 2008）、学龄儿童（e.g., Huizinga et al., 2006; Lehto et al., 2003）、青少年（e.g., Feng et al., 2022; Friedman et al., 2011; Xu et al., 2013）和成人（e.g., Feng et al., 2022; Friedman & Miyake, 2004; Miyake et al., 2000, 2001）等群体中执行功能不同成分的发展变化。

与潜在变量分析不同的是，网络分析是考虑所有变量之间的共同性之后如何关联的。通过数据驱动方法，网络分析可为识别独特和共同的认知机制提供新的见解，常用于理解认知结构，特别是复杂的相互关联的成分，如智力、精神病理学和人格等。Younger 等人（2023）首次将该方法用于分析执行功能结构，发现至少到 10 岁（五年级）时加利福尼亚儿童在 8 项任务上的表现才可识别出三个稳定的成分，并且这些成分的分离至少要持续到 14 岁（八年级）。该方法从整体角度检验执行功能的结构，而非构建和比较多个竞争模型，更能反映数据的组织特点，可作为潜在变量分析的补充（蔡玉清 等，2019; Younger et al., 2023）。

### 1.3 执行功能的结构——统一性和多样性

执行功能在个体生命早期就已经出现，并持续发展到青春期中后期或成年早期（Huizinga et al., 2006; McAuley et al., 2010）；其结构从学前阶段开始逐渐分化（Best & Miller, 2010），从单一结构逐渐转变为当前普遍接受的一系列多样的、互动的加工过程（Karr et al., 2018）。Garon 等人（2008）基于认知技能发展的连续性和层次性，提出了关于执行功能的整合模型（The Integrative Executive Functions Model），认为执行功能各成分是基于更为简单的认知技能（Simpler Cognitive Skills）发展而来，或是更为简单的认知技能相互协调的结果。这些更为简单的认知技能对教育发展至关重要，如目标保持等，是有效执行各项任务的必要条件；而组合技能（Combination Skills）是与执行功能的实际组成成分或中央执行相关的技能。并且，这些技能随着个体脑发育逐渐变得明显、成熟。该模型强调了执行功能统一性不可忽视的作用。在儿童和成人群体中，不同的执行功能任务测验是相关关联的；更为重要的是，在不同的年龄组中所有潜在因素均存在相关性，表明存在一个共同的过程（Friedman & Miyake, 2004; Hughes & Devine, 2019; Lehto et al., 2003; Miyake et al., 2000）。

执行功能的统一性/多样性框架（Unity/Diversity Framework; Miyake & Friedman, 2012）认为，每个执行功能（如，工作记忆）均可分解为所有任务的普遍成分或统一性（对应着一般因子），以及独特成分或多样性（对应着特异因子，如工作记忆特异因子）。通过验证性因子分析方法，执行功能各成分相关的任务可更为纯净地映射到潜在的认知过程（包括执行功能一般因子和特异因子）上。Friedman 及其合作者认为，双因子模型的一般因子可表示执行功能的统一性，并且该模型可更为直接、清晰地表征统一性与多样性之间的关系（Friedman & Miyake, 2017; Friedman et al., 2008）。然而，Karr 等人（2018）汇总了采用验证性因子分析探讨执行功能的 40 项研究，对 46 个样本（$N = 9{,}756$）的结果进行重新分析却发现，执行功能的最佳模型在不同发展阶段并不同，学前儿童更多支持单一因子或两因子模型，学龄儿童更多支持三因子模型，青少年/成人则支持三因子或双因子模型。也就是说，学前和学龄儿童中可能并不存在双因子模型，无法提取出双因子模型中的一般因子和特异因子。为了更好地描绘童年中期儿童执行功能的统一性与多样性，当双因子模型无法与行为数据实现较好拟合时，本文认为可以提取单因子模型中的一般因子作为统一性的指标，而其它与执行功能成分直接相关的因子作为多样性的指标。据我们所知，尚未有研究直接考察中国教育文化背景下童年中期儿童执行功能的结构，无法确定该阶段中国儿童执行功能的结构是否与西方群体类似。

### 1.4 本研究目的

通过梳理国内外执行功能相关主题的进展，现有研究亟待探究的问题如下：一方面，执行功能的衡量指标不纯粹，国内相关研究普遍采用单个任务衡量执行功能的成分，受到任务杂质的影响，无法提取“干净”成分；另一方面，缺乏关注中国文化背景下童年中期儿童执行功能的结构及稳定性。虽然，研究者描绘、探究了执行功能各成分（主要是抑制控制、工作记忆和认知灵活性）的发展趋势，对于理解不同阶段的认知发展特征具有重要参考价值。例如，齐玥等人（2021）概述了执行功能不同成分的发展特点。但执行功能的结构从学前阶段开始逐渐分化（Best & Miller, 2010）；有关执行功能的个体差异研究也发现，执行功能存在统一性和多样性（Miyake & Friedman, 2012; Miyake et al., 2000），并且在不同年龄阶段的结构不同（Karr et al., 2018）。例如，Brydges 等人（2014）发现童年中后期执行功能存在单因子模型和两因子模型。因此，对于童年中期儿童执行功能统一性和多样性有待深入探究。

通过对当前执行功能发展研究进行回顾，可发现存在以下局限：（1）不同的研究采用不同的任务评估了相同成分。例如，评估抑制控制时，采用 Go/Nogo 或 Flanker 等任务；（2）采用单个任务衡量某个成分，无法提取有效的“干净”成分；（3）多涉及横断设计，仅有限数量的纵向研究（e.g., Brydge et al., 2014; Friedman et al., 2011, 2016），并且忽略了执行功能的纵向不变性，无法确定哪些因子结构稳定且可重复；（4）多数研究的样本量较小，一般为 100 ~ 300（e.g., Brydge et al., 2014; $N = 135$），无法获得稳定的效果量，并且潜在变量分析的参数较多，需要更大的样本量来获得稳定的结构；（5）缺乏衡量童年中期儿童执行功能统一性和多样性的有效指标，两者并不是非此即彼，在儿童发展中均发挥重要作用，无法同时考察两者的贡献性。

鉴于上述问题和局限性，本文结合执行功能有关统一性/多样性的两个模型，从行为学层面，采用间隔一年的纵向设计，通过基于表现的多任务测试评估一个较大儿童样本的执行功能，并通过潜在变量分析和网络分析，确定中国文化背景

下童年中期儿童执行功能的结构及稳定性。本研究选取追踪起始年级为小学三年级，主要基于以下考虑：（1）儿童期和青少年期是身心发展的高速阶段，该阶段也是认知控制和情绪发展的重要阶段，认知控制作为高级心理功能，其正常发展能够确保个体具有充分的主观能动性和自我控制性，并对其他系统起到相互制约和促进的作用（齐玥 等，2021）；（2）该阶段儿童的执行功能不同成分已获得进一步发展（Lehto et al., 2003），这方面的相关研究将注重对执行功能风险因素的规避和保护性因素的提升，对认识和训练执行功能更具参考价值。

## 2 方法

### 2.1 被试

本研究的数据选自一项正在进行的为期 2 年的纵向研究项目，该项目主要关注领域一般认知能力是否以及如何在童年中期儿童数学焦虑的发展中发挥作用。该项目从济南市三所公立小学招募三至四年级参与者。在 2021 年秋季学期到 2023 年春季学期进行数据采集，每波间隔 6 个月。本研究选取第一次和第三次的数据进行分析。该项目获得当地大学伦理委员会批准。每次调查前均获得家长和学校的知情同意。

第一次共 1192 名儿童参加（98.07%完成了三项以上的任务并达到合格）；由于疫情防控原因，第三次仅 956 名儿童参与（91.53%完成了三项以上的任务并达到合格），两次共计 1383 名儿童参与。由于转学（包括转出和新转入）、请假、退学等客观因素，并非所有的儿童均参与所有时间点的测查，如果儿童在两个时间点中均没有提供任何关于任务的合格数据，则被排除在分析之外（$N = 510$）；如果儿童在两个时间点中任意一个时间点未有效完成所有任务的一半，则被排除在分析之外（$N = 105$）；剔除瑞文智力测验等级为五的儿童（$N = 12$）。最终参与数据分析的有效被试数量为 756 名，其中 312 名在所有时间点的所有任务上均提供了有效数据。具有完整数据的被试与具有不完整数据的儿童在利手（$\chi^2 = 0.32, p > 0.05$）上差异不显著，在年龄（$t = 3.07, p = 0.002$）、性别（$\chi^2 = 6.48, p = 0.011$）和主要变量上差异显著（$|t|s \geq 2.07, ps < 0.05$），具有完整数据的儿童在所有时间点上表现更佳。Little’s MCAR 检验结果显示显著（$p < 0.001$），表明数据缺失不是完全随机缺失。最终参与分析的 756 名儿童的基本信息见表 1，平均年龄为 9.25，76.04%父亲和 70.60%母亲接受高中以上教育。

表 1 参与者的人口学信息

| 变量 | | 数据 |
|---|---|---|
| **儿童特征** | | $N = 756$ |
| 年龄（岁） | | |
| | 范围 | 8.17 ~ 10.75 |
| | $M$ ($SD$) | 9.25 (0.56) |
| 性别（男，%） | | 364 (48.15%) |
| 民族（汉族，%） | | 750 (99.21%) |
| 利手（右利手，%） | | 746 (98.68%) |
| IQ（%） | | |
| 5 ~ 25 | | 201 (26.59%) |
| 25 ~ 75 | | 420 (55.56%) |
| 75 ~ 95 | | 124 (16.40%) |
| ≥ 95 | | 11 (1.45%) |
| **家庭特征** | | |
| 父母受教育程度（父亲/母亲, %） | | $N = 722$ / 721 |
| 小学及小学以下 | | 1.39% / 2.08% |

| | |
|---|---|
| 初中（含初中未毕业） | 22.58% / 27.32% |
| 高中或中专（含高中未毕业） | 29.92% / 31.21% |
| 大专（含夜大、电大、党校） | 27.01% / 21.64% |
| 大学本科 | 18.14% / 17.06% |
| 研究生（硕士或博士） | 0.97% / 0.69% |

## 2.2 工具

### 2.2.1 执行功能测验

采用行为测试集评估执行功能。该测试集包含 6 项任务，分别为两项工作记忆任务、两项抑制控制任务和两项认知灵活性任务。所有任务在计算机上运行，刺激材料呈现在 15 寸 Lenovo 显示器上（分辨率 1,024 × 68，刷新率为 60 Hz），屏幕背景色为灰色，前景色为白色。字体均为 Times New Roman，字号为 32。

**（1）抑制控制**

参考 Maasalo 等人（2021）的研究，采用修订版的、对于儿童更为友好的 Go/Nogo 任务和 Oddball 任务（Downes et al., 2017; Seer et al., 2016）评估儿童的抑制控制能力，测量个体抑制优势反应的能力。

*Oddball 任务*。在屏幕的中央呈现不同大小的黑豆（大黑豆的视角约为 1.15° × 2.29°；小黑豆的视角约为 0.46° × 0.92°），要求儿童看到小黑豆（150 个试次，占总试次的 75%）时尽可能快而准确地用左侧食指按“F”键，看到大黑豆（50 个试次， 占总试次的 25%）时则用右侧食指按“J”键。记录儿童按键的正确率和反应时间。大黑豆试次为目标试次，计算击中率（即目标试次正确按键）、错误警觉（即非目标试次错误按键）和辨别力（$d' = Z_{击中率} - Z_{错误警觉}$）。其中辨别力 d'作为儿童抑制控制能力的行为指标，得分越高表示儿童的辩别力越好。

*Go/Nogo 任务*。在屏幕的中央呈现不同颜色的点（视角约为 1.83° × 1.83°），因颜色不同被分为 Go 刺激和 Nogo 刺激。要求儿童看到 Go 刺激（150 个试次，占总试次的 75%）时尽可能快而准确地按“F”键进行反应，看到 Nogo 刺激（50 个试次，占总试次的 25%）时不反应。记录儿童按键的正确性和反应时间，Nogo 试次为目标试次，计算击中率、错误警觉和辨别力 d'。其中辨别力 d'作为儿童抑制控制能力的行为指标。

上述两个任务的流程参数相同，即呈现注视点 500 ms，然后呈现刺激 500 ms，紧接着是 1000 ms 的空屏；试次间隔为 800 ms 到 1000 ms，以减少儿童预期（详细流程见图 1A 和图 1B）。要求儿童在刺激呈现屏和随后的空屏（共 1500 ms）进行反应。正式实验之前，儿童需完成 10 个练习试次以确保理解实验流程，正确率达到 90%方可进入正式实验。每个任务包含 4 个 Blocks，每个 Block 包含 50 个试次，Block 之间有短暂的休息。每个任务持续时间约为 5 min。

**（2）工作记忆**

参考 Maasalo 等人（2021）的研究，采用 N-back 任务中的 1-back 任务（Pelegrina et al., 2020; Yaple & Arsalidou, 2018）和韦氏智力测验中的数字广度子测验（Digital Span; Wechsler, 2002）评估儿童的工作记忆容量。

*1-back 任务*。采用数字 1-back 任务，刺激是 9 个阿拉伯数字（1 ~ 9）。在每个试次中，屏幕中央呈现一个数字，持续时间为 400 ms；之后呈现 1000 ms 的空屏，要求儿童在刺激呈现屏和空屏（共 1400 ms）进行按键反应；试次间隔为 500 ms（详细流程见图 1C）。儿童需要判断看到的数字与前一个数字是否相同，若相同按“F”键，不同按“J”键。正式实验之前，儿童需完成 10 个练习试次，以确保理解实验流程，正确率达到 90%方可进入正式实验。分为 4 个 Blocks，每个 Block 包含 50 个试次，每个 Block 之间有短暂的休息。数字相同试次为目标试次，

占总试次的 25%。该任务持续时间为 5 min。记录儿童按键的正确性和反应时间，计算击中率、错误警觉、辨别力 d'，其中辨别力 d'作为儿童工作记忆能力的行为指标。

*倒序数字广度任务*（*Backward Digital Span*）。给儿童呈现一串数字（数字集合），数字范围为 1 ~ 9。数字集合的长度为 2 ~ 6，每类长度的数字集合包含 6 个。每个数字集合中的数字以每秒一个的速度播放，按数字集合长度递增的方式依次向被试播放数字集合。每个数字集合播放完毕后，要求儿童按倒序的方式复述耳机中播放的数字。每类长度的数字集合连续正确复述 4 个进入下一类长度的数字集合，相同长度的数字集合连续错误 3 个则停止该测验。记录儿童复述正确的个数并作为数字广度的指标，得分越高表示儿童的数字工作记忆能力越好。

**（3）认知灵活性**

参考 Maasalo 等人（2021）的研究，采用计算机版的 WCST（Lange et al., 2017）和 TMT（Takacs & Kassai, 2019）任务评估儿童的认知灵活性。

*WCST*。根据形状（正方形、圆形、五角星和十字形）、颜色（红色、绿色、黄色和蓝色）和数量（1、2、3 和 4）制作了 64 张卡片（卡片视角约为 5.72° × 5.72°），其中 4 张为目标卡片。为保证被试分类规则的确定性，要求测试卡片与每一张目标卡片共有的属性不超过一个，最后筛选出 24 张卡片作为测试卡片（Lange et al., 2017）。屏幕上呈现 4 张目标卡片和 1 张测试卡片，其中目标卡片呈现在测试卡片正上方，目标卡片自左向右对应数字 1 ~ 4。要求儿童根据三种可能的分类规则之一匹配目标卡片和测试卡片，并根据目标卡片对应的数字进行按键反应。每个试次中，先在屏幕中央呈现 800 ms 的注视点，然后呈现刺激界面，按键消失；间隔 800 ms 后呈现反馈，持续时间为 400 ms，告知被试分类规则是否正确（同时也是提示下一个试次中的分类规则，继续保持上一个试次运用的分类规则，还是找寻新的分类规则）（详细流程见图 1D）。根据反馈可将试次分为转换试次和保持试次，为减少儿童预期，以不可预测的方式改变规则，即当保持试次连续正确的个数为 4 ~ 6 时进行随机变换新的规则。三种分类规则以乱序的方式随机呈现两次，儿童共需完成至少 24 个试次、最多 36 个试次。该任务持续时间约为 10 min，记录转换试次和保持试次的正确性和反应时间，计算儿童分类的总正确率作为认知灵活性的行为指标。

*TMT*。该测验为一项计时测验，可用于评估儿童的认知灵活性能力，分为 2 个部分。在 A 部分中，包含编号为 1 ~ 25 的圆圈，要求儿童从编号为 1 的圆圈开始，以升序的方式通过点击鼠标将所有的圆圈连接起来（如，1−2−3−...）；在 B 部分中，同样包含 25 个圆圈，但一半圆圈为数字 1 ~ 13 编号，另一半圆圈为字母 A ~ L 编号，要求被试以升序、交替的方式通过画线将所有的圆圈连接起来（如，1−A−2−B−3−C−......）。该任务中 A 部分持续时间为 2.5min，B 部分为 5 min，记录每个部分完成的实际所需时长。计算 B 部分与 A 部分的时长差值（即转换代价），差值越小表明儿童的认知灵活性越好。

需要说明的是，在 1-back、Go/Nogo、Oddball 任务中，对目标刺激的正确反应被定义为击中，对非目标刺激的错误反应定义为错误警觉。三个任务中，正式实验剔除无效反应或 RT < 200 ms 的试次后，有效率需达到 80%以上；对非目标刺激正确反应的试次需占有效试次的 50%以上。将未满足上述两个条件的儿童数据视为缺失（Huizinga et al., 2006）。本研究中，执行功能测验所有任务指标的重测信度 ICC 分别为 1-back：0.56；Digit Span：0.44；Oddball：0.21；Go/Nogo：0.13；WCST：0.35；TMT：0.18。

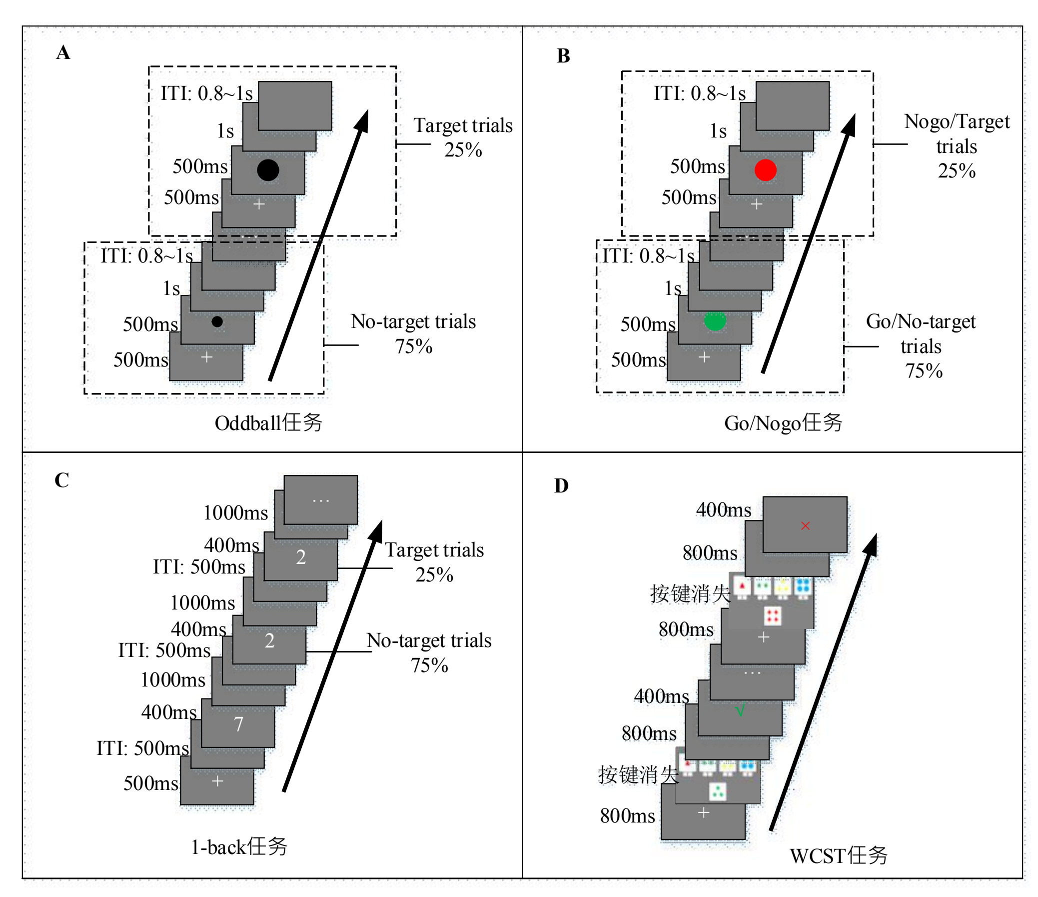



图 1 执行功能任务测试集中部分任务的流程

### 2.2.2 智力

采用 Raven 等人（1983）开发、并由张厚粲和王晓平（1989）修订的瑞文标准化推理测验（Raven's Standard Progressive Matrices, SPM），评估儿童的观察力及清晰思维的能力。该测验是一项纯粹的非文字智力测验，广泛应用于无国界的智力/推理能力测查，可评估个体的非言语智力。该测验分为 5 个单元，共 60 张图片，每张图片为一个渐进矩阵构图，矩阵的结构越来越复杂，从一个层次到多个层次的演变，要求的思维操作也是从直接观察到间接抽象推理的渐进过程。答对 1 题得 1 分，总分 60 。根据儿童的年龄和常模，将原始分转换为标准等级分，位于等级五的视为智力缺陷，需要被排除。

### 2.2.3 人口学变量

儿童的性别、年龄、民族和家庭社会经济地位（主要为父亲/母亲受教育水平）由父母报告，儿童利手情况由儿童自我报告。

## 2.3 研究流程

在三所公立小学中招募儿童参加“冷”执行功能的测查。在收集数据之前，获得学校和班主任的知情同意，之后由儿童给家长一份包含知情同意信息的家长问卷，家长自愿签署是否同意让孩子参加研究，并完成相关家长方面的调查。在学校机房中，由 4 名训练有素的心理学研究生担任研究助理，负责数据收集工作，历时 60 min。计算机任务测试阶段需要每位儿童完成执行功能测试集，共包含 6 项任务。对于每项任务，研究助理指导儿童理解任务目标，并进行练习。每项任务中，练习试次的正确率达到 90%以上方可进入正式试验。为感谢儿童的参与，每次完成任务的儿童将获得一份精美学习用品。

## 2.4 数据管理与分析

采用 IBM SPSS 26.0（IBM Corp., Armonk, N.Y., USA）对问卷测查阶段的数据进行双录，使用 Excel 2010 中的条件格式功能对数据的准确性进行检验；对于计算机任务测查阶段的数据，采用 Excel 2010 中的分类汇总功能和 vlookup 函数计算每个儿童每项任务的指标。在 IBM SPSS 26.0、JASP 0.16.4.0（JASP Team, 2022）、Mplus 8.3（Muthén & Muthén, 1998−2017, CA）和 R 4.1.2（R Core Team, 2020）中对所有数据进行整理与分析，包含以下过程：

首先，在 IBM SPSS 26.0 中，对计算机任务测查阶段的数据进行清理和转换。为方便理解，将数据进行转换，使得分越高表示越好的表现。因此，对于 TMT 任务中的转换代价，本研究将其乘以−1，转换为负值（即−RT）（Brydges et al., 2014; Feng et al., 2022）。采用期望最大化算法（Expectation-Maximization Algorithm, EM）分别对每个时间点的缺失数据单独进行插补（宋枝璘 等，2022）。

其次，在 JASP 0.16.4.0 中，进行描述统计分析和相关分析，提供执行功能和数学焦虑的一般性特征；在 Mplus 8.3 中，通过验证性因子分析，构建并验证 12 种执行功能潜在因子结构，采用极大似然（Maximum Likelihood, ML）估计模型。模型拟合指标包含卡方检验（Chi-Square Test, $\chi2$）、比较性配适指标（Comparative Fit Index, CFI）、塔克-刘易斯指数（Tucker-Lewis Index, TLI）、近似均方根误差（Root-Mean Square Error of Approximation, RMSEA）和标准化均方根残差（Standardized Root-Mean-Square Residual, SRMR）。模型拟合的标准如下：CFT 与 TLI 大于 0.90，RMSEA 小于 0.05，SRMR 小于 0.08（Hu & Bentler, 1999）；通过将拟合较好模型的 AIC 和 BIC 指数进行比较，选择与行为数据拟合最佳的模型作为执行功能的结构（邱皓政，林碧芳，2019）；

再次，在 R 4.1.2 中估计了 6 项执行功能任务之间偏相关系数的高斯图形模型（GGM）（Epskamp, 2015）。应用最小绝对收缩和选择算子（LASSO）进行正则化，谨慎地识别网络边缘，从而准确地揭示底层网络。为了直观比较不同网络，用两个网络中节点的平均位置统一了网络布局，并将网络中最小的边缘值设为 0.04。此外，进一步通过 EGAnet 包（Younger et al., 2023）检测不同任务是否在网络内形成了一个或多个社区（community）。

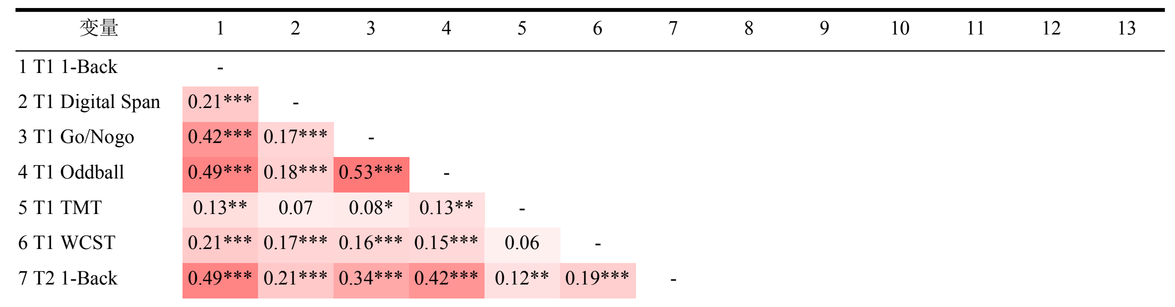
最后，在 Mplus 8.3 中对执行功能的统一性和多样性进一步开展纵向等值性检验。

## 3 结果

### 3.1 初步分析

首先，对执行功能任务相关指标进行描述统计和相关分析（见表 1）。结果发现，本研究中执行功能相关任务指标呈正态或近似正态分布（$-0.62 \le Skewness \le 0.53$）。同一个时间点中执行功能各任务表现之间存在显著的正相关（T1: $rs \ge 0.10$, $ps < 0.05$; T3: $rs \ge 0.09$, $ps < 0.05$）；执行功能相同任务在不同时间点之间存在显著的正相关（$rs \ge 0.17$, $ps < 0.001$）。

表 1 变量的描述统计和相关结果



| 变量 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 T1 1-Back | - | | | | | | | | | | | | |
| 2 T1 Digital Span | 0.21*** | - | | | | | | | | | | | |
| 3 T1 Go/Nogo | 0.42*** | 0.17*** | - | | | | | | | | | | |
| 4 T1 Oddball | 0.49*** | 0.18*** | 0.53*** | - | | | | | | | | | |
| 5 T1 TMT | 0.13** | 0.07 | 0.08* | 0.13** | - | | | | | | | | |
| 6 T1 WCST | 0.21*** | 0.17*** | 0.16*** | 0.15*** | 0.06 | - | | | | | | | |
| 7 T2 1-Back | 0.49*** | 0.21*** | 0.34*** | 0.42*** | 0.12** | 0.19*** | - | | | | | | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 T2 Digital Span | 0.24*** | 0.38*** | 0.24*** | 0.30*** | 0.04 | 0.15*** | 0.26*** | - | | | | | |
| 9 T2 Go/Nogo | 0.24*** | 0.13*** | 0.33*** | 0.30*** | 0.08* | 0.11** | 0.33*** | 0.29*** | - | | | | |
| 10 T2 Oddball | 0.26*** | 0.10* | 0.29*** | 0.42*** | 0.07 | 0.02 | 0.47*** | 0.22*** | 0.47*** | - | | | |
| 11 T2 TMT | 0.05 | 0.13** | 0.11** | 0.10* | 0.17*** | 0.13*** | 0.14*** | 0.13*** | 0.06 | 0.09* | - | | |
| 12 T2 WCST | 0.19*** | 0.10* | 0.18*** | 0.16*** | 0.10* | 0.30*** | 0.23*** | 0.15*** | 0.18*** | 0.15*** | 0.10** | - | |
| 13 年龄 | −0.04 | 0.05 | 0.04 | 0.02 | 0.05 | 0.05 | 0.12** | 0.16*** | 0.04 | 0.07 | 0.08* | 0.08* | - |
| 14 性别 | 0.15*** | 0.03 | 0.12** | 0.22*** | −0.07 | 0.06 | 0.23*** | 0.06 | 0.08* | 0.22*** | 0.01 | 0.06 | −0.03 |
| *N* | 582 | 661 | 751 | 689 | 694 | 755 | 717 | 718 | 752 | 735 | 731 | 754 | 756 |
| Mean | 1.19 | 18.38 | 2.92 | 2.22 | −1.18 | 0.51 | 1.91 | 20.35 | 3.41 | 2.63 | −0.85 | 0.56 | – |
| *SD* | 0.85 | 5.65 | 1.06 | 0.99 | 0.6 | 0.17 | 0.92 | 5.73 | 1.18 | 1.09 | 0.41 | 0.18 | – |
| Skewness | 0.53 | 0.25 | 0.08 | 0.25 | −0.62 | 0.05 | 0.11 | −0.09 | −0.01 | 0.17 | −0.52 | −0.51 | −0.07 |
| Min | 0.01 | 9 | 0.02 | 0.01 | −3.19 | 0.13 | 0.02 | 9 | 0.1 | 0.12 | −2.28 | 0.09 | 0 |
| Max | 3.72 | 30 | 6.18 | 5.48 | 0.64 | 0.89 | 5.05 | 30 | 6.18 | 6.18 | 0.43 | 0.89 | 1 |

注: 1-back 和 Digital Span 为工作记忆相关任务；Go/Nogo 和 Oddball 为抑制控制相关任务；TMT 和 WCST 为认知灵活性相关任务。红色表示正相关，蓝色表示负相关，颜色越深表示相关越大。* $p < 0.05$，** $p < 0.01$，*** $p < 0.001$。

### 3.2 不同时间点执行功能结构的候选模型

为探索童年中期儿童执行功能的结构，采用验证性因子分析，在两个时间点上分别构建执行功能的 12 种候选模型（又称竞争模型，包括相关因子模型和双因子模型），模型拟合结果见表 2。相关因子模型中的“G”、“I + W/C”、“I/C + W”和“I + W + C”模型，以及双因子模型中的“G + W”、“G + C+ W”、“G + I + W”和“G + I + C”模型，在所有时间点模型拟合均较好，而其余模型在至少一个时间点上拟合较差或无法收敛。但在双因子模型中，拟合较好的模型的某些任务指标在两个时间点的抑制控制特异因子或工作记忆特异因子或认知灵活特异因子上的载荷均不显著（*ps* > 0.05），因此，最终保留相关因子模型中的“G”、“I/W + C”、“I + W/C”、“I/C + W”和“I + W + C”模型。与早期的研究一致（Friedman et al., 2008; Miyake et al., 2000），“I + W + C”模型表明了执行功能的多样性，并且三个因子具有较高的相关性（*ps* < 0.001），“G”模型表明了执行功能的统一性；此外，“I/W + C”、“I + W/C”模型和“I/C + W”模型拟合也较好，支持了抑制控制、认知灵活性及工作记忆两两之间是非独立的观点。

表 2 两个时间点上执行功能相关因子候选模型的拟合指数

| 模型 | 时间点 | $\chi^2$ | *df* | *p* | CFI | TLI | AIC | BIC | *a*BIC | RMSEA (90%CI) | SRMR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 相关因子模型——统一性 | | | | | | | | | | | |
| 1. G | T1 | 16.927 | 9 | 0.050 | 0.983 | 0.972 | 9988.392 | 10071.697 | 10014.539 | 0.034 [0.001, 0.059] | 0.026 |
| | T2 | 17.742 | 8 | 0.023 | 0.981 | 0.964 | 10897.708 | 10985.641 | 10925.308 | 0.040 [0.014, 0.066] | 0.026 |
| 相关因子模型——多样性 | | | | | | | | | | | |
| 2. I/W + C | T1 | 16.745 | 8 | 0.033 | 0.981 | 0.965 | 9990.140 | 10078.073 | 10017.740 | 0.038 [0.010, 0.064] | 0.026 |
| | T2 | 14.981 | 7 | 0.036 | 0.984 | 0.966 | 10897.031 | 10989.591 | 10926.083 | 0.039 [0.009, 0.066] | 0.024 |
| 3. I + W/C | T1 | 7.261 | 8 | 0.509 | 1.000 | 1.003 | 9980.826 | 10068.759 | 10008.426 | 0.000 [0.000, 0.040] | 0.018 |
| | T2 | 9.044 | 7 | 0.250 | 0.996 | 0.991 | 10890.998 | 10983.558 | 10920.050 | 0.020 [0.000, 0.052] | 0.017 |
| 4. I/C + W | T1 | 16.216 | 8 | 0.039 | 0.983 | 0.967 | 9989.708 | 10077.641 | 10017.308 | 0.037 [0.008, 0.063] | 0.026 |
| | T2 | 15.358 | 7 | 0.032 | 0.984 | 0.965 | 10897.503 | 10990.064 | 10926.556 | 0.040 [0.011, 0.091] | 0.024 |
| 5. I + W + C | T1 | 5.083 | 6 | 0.533 | 1.000 | 1.005 | 10003.161 | 10100.350 | 10033.666 | 0.000 [0.000, 0.043] | 0.015 |
| | T2 | 19.799 | 6 | 0.003 | 0.973 | 0.932 | 10902.972 | 11000.161 | 10933.478 | 0.055 [0.029, 0.083] | 0.021 |
| 双因子模型 | | | | | | | | | | | |
| 1. G + I | T1 | 6.353 | 7 | 0.4992 | 1.000 | 1.003 | 9982.826 | 10075.387 | 10011.879 | 0.000 [0.000, 0.042] | 0.018 |
| | T2 | | | | | | No Convergence | | | | |
| 2. G + W | T1 | 14.189 | 7 | 0.0479 | 0.985 | 0.967 | 9991.708 | 10084.269 | 10020.760 | 0.037 [0.003, 0.065] | 0.026 |
| | T2 | 36.404 | 7 | <0.001 | 0.942 | 0.876 | 10921.486 | 11014.047 | 10950.538 | 0.075 [0.052, 0.099] | 0.033 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3. G + C | T1 | 14.652 | 7 | 0.0407 | 0.984 | 0.965 | 9992.140 | 10084.701 | 10021.193 | 0.038 [0.007, 0.065] | 0.026 |
| | T2 | No Convergence | | | | | | | | | |
| 4. G + I + W | T1 | 0.964 | 4 | 0.9151 | 1.000 | 1.024 | 9982.427 | 10088.871 | 10015.837 | 0.000 [0.000, 0.020] | 0.005 |
| | T2 | 15.244 | 5 | 0.6639 | 0.898 | 0.94 | 10906.492 | 11008.309 | 10938.45 | 0.052 [0.023, 0.083] | 0.023 |
| 5. G + C + W | T1 | 0.645 | 4 | 0.958 | 1.000 | 1.027 | 9982.155 | 10088.600 | 10015.565 | 0.000 [0.000, 0.000] | 0.004 |
| | T2 | 18.889 | 4 | 0.0008 | 0.971 | 0.890 | 10905.529 | 11011.974 | 10938.940 | 0.070 [0.040, 0.103] | 0.020 |
| 6. G + I + C | T1 | 4.813 | 4 | 0.3070 | 0.998 | 0.994 | 9986.148 | 10092.593 | 10019.558 | 0.016 [0.000, 0.059] | 0.014 |
| | T2 | 18.084 | 4 | 0.0012 | 0.972 | 0.896 | 10903.946 | 11010.391 | 10937.356 | 0.068 [0.038, 0.102] | 0.020 |
| 7. G + I + C + W | T1 | No Convergence | | | | | | | | | |
| | T2 | No Convergence | | | | | | | | | |

注：相关因子模型中 W 为工作记忆相关因子，I 为抑制控制相关因子，C 为认知灵活性相关因子，G 为一般因子，I/C 为抑制控制与认知灵活性共同因子，I/W 为抑制控制与工作记忆共同因子，W/C 为工作记忆与认知灵活性共同因子。

对 5 种拟合较好的竞争性相关因子模型进行比较，以确定该阶段不同时间点儿童的最佳拟合模型（结果见表 3）。模型比较显示，两个时间点中，与“G”、“I/W + C”、“I/C + W”和“I + W + C”模型相比，“I + W/C”模型中 AIC、BIC 更小，这表明童年中期儿童的执行功能结构主要以两因子“I + W/C”模型存在。

**表 3 拟合较好的相关因子模型比较**

| 模型比较 | T1 | | T2 | |
|---|---|---|---|---|
| | ΔAIC | ΔBIC | ΔAIC | ΔBIC |
| “G” VS. “I/W + C” | -1.748 | -6.376 | 0.677 | -3.950 |
| “G” VS. “I + W/C” | 7.566 | 2.938 | 6.710 | 2.083 |
| “G” VS. “I/C + W” | -1.316 | -5.944 | 0.205 | -4.423 |
| “G” VS. “I + W + C” | -14.769 | -28.653 | -5.264 | -14.52 |
| “I/W + C” VS. “I + W/C” | 9.314 | 9.314 | 6.033 | 6.033 |
| “I/W + C” VS. “I/C + W” | 0.432 | 0.432 | -0.472 | -0.473 |
| “I/W + C” VS. “I + W + C” | -13.021 | -22.277 | -5.941 | -10.57 |
| “I + W/C” VS. “I/C + W” | -8.882 | -8.882 | -6.505 | -6.506 |
| “I + W/C” VS. “I + W + C” | -22.335 | -31.591 | -11.974 | -16.603 |
| “I/C + W” VS. “I + W + C” | -13.453 | -22.709 | -5.469 | -10.097 |

### 3.3 不同时间点执行功能相关任务的网络特征

接下来，分别对两个时间点上执行功能的 6 项任务指标进行网络分析，以了解不同任务是如何关联的，结果见图 2。整体上，两个时间点的网络模型相似，T2 时间点上网络比 T1 时间点的更为复杂性；关于网络的组块，EGAnet 包中的 Spinglass 算法显示两个时间点上均只有一个社团。该结果强调在三至四年级儿童发展中，执行功能的结果变化不大，为单一成分，但这些任务之间的关系逐渐复杂化，即执行功能随时间推移不断完善。因此，非常有必要重视该阶段执行功能的统一性。

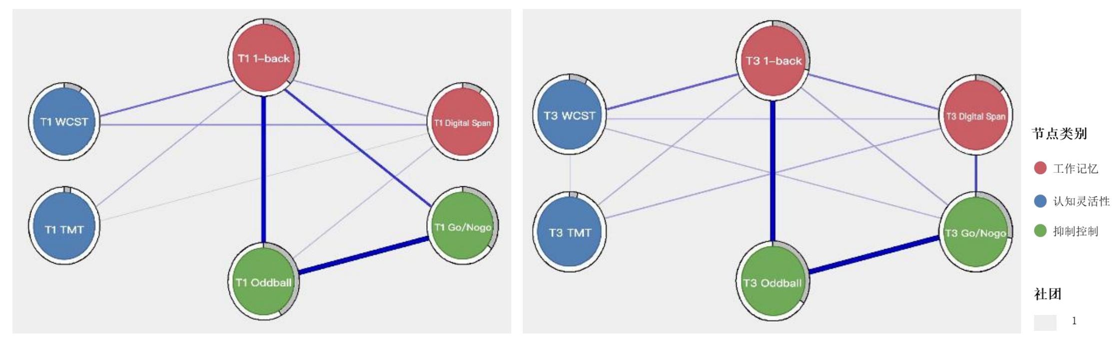


图 2　执行功能在两个时间点上的网络特征

注：节点之间的连线表示任务之间的偏相关，线越粗表示偏相关系数越大。为了直观比较不同网络，通过两个网络中节点的平均位置统一了网络布局。节点周围圆环中的彩色区域描绘了可预测性。

### 3.4 执行功能统一性和多样性的纵向等值性检验

结合上述潜在变量分析和网络分析的结果，本研究认为童年中期儿童执行功能的统一性和多样性均不可忽略。然而，本研究中双因子模型拟合较差，为衡量执行功能的统一性和多样性，提取单因子"G"模型中的一般因子作为执行功能统一性指标，提取执行功能最优模型——两因子"I + W/C"模型中与执行功能成分直接相关的因子作为执行功能多样性的指标。两个时间点上执行功能的统一性和多样性因子载荷情况见图 3。

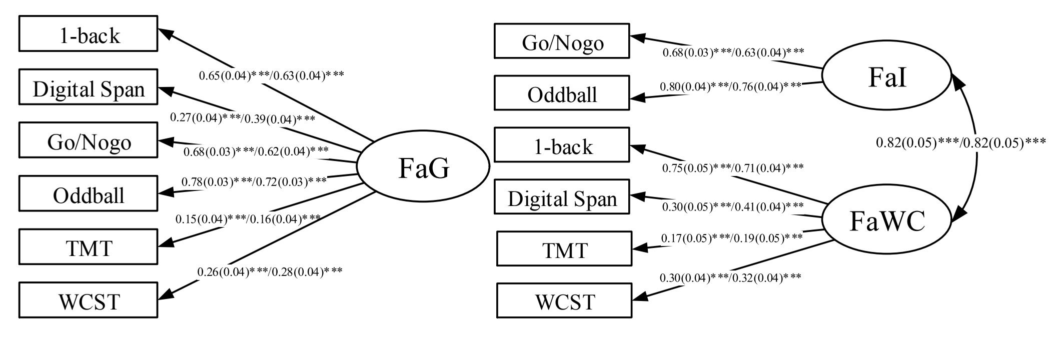



图 3 两个时间点上执行功能的统一性和多样性

注：左图为"G"模型，FaG 为"G"模型中的一般因子；左图为"I+ W/C"模型，FaI 和 FaWC 为"I+ W/C"模型中的成分因子，分别对应一致控制因子和工作记忆与认知灵活性共同因子；斜线左侧为 T1 时间点标准化系数，斜线右侧为 T2 时间点标准化系数；括号内为标准误。*** $p < 0.001$。

运用多组验证性因子分析（Multiple Groups Confirmatory Factor Analysis），建构一系列嵌套模型，分别对执行功能统一性和多样性和数学焦虑进行纵向等值性检验。当完全等值性无法满足时，采用 Forward 方法（Jung & Yoon, 2016）进行部分等值性设置。模型拟合的标准如下：CFT 与 TLI 大于 0.90，RMSEA 小于 0.05，SRMR 小于 0.08 （Hu & Bentler, 1999）。嵌套模型比较基于拟合指数$\Delta\chi^2$、$\Delta$CFI 和$\Delta$RMSEA 的变化进行评估，即当$\Delta\chi^2$未达到显著水平，表明限制更多的模型未导致模型拟合变差，模型等值可接受；当 CFI 的变化小于 0.010、RSMEA 的变化小于 0.015 时，表明限制更多的模型未导致模型拟合变差，模型等值可接受（Chen, 2007; Satorra & Bentler, 2001）。由于卡方检验受样本量影响较大（Feng et al., 2022），更依赖于后者进行判断。纵向等值性检验结果（见表 4）表明，执行功能统一性（即"G"模型中的一般因子）和多样性（即"I + W/C"模型中的成分相关因子）均满足完全单位等值。

**表 4 执行功能统一性和多样性的纵向等值性检验**

| 模型 | $\chi^2/df$ | $p$ | CFI | TLI | AIC | BIC | $a$BIC | RMSEA(90%CI) | SRMR | $\Delta\chi^2/\Delta df$ | ΔCFI | ΔRMSEA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | "G" 模型 | | | | | | |
| 完全形态等值 | 2.462 | <0.001 | 0.953 | 0.933 | 20481.605 | 20680.610 | 20544.068 | 0.044[0.034,0.054] | 0.039 | – | – | – |
| **完全单位等值** | **2.498** | **<0.001** | **0.946** | **0.932** | **20485.905** | **20661.771** | **20541.105** | **0.045[0.035,0.054]** | **0.045** | 2.835 | **−0.007** | **0.001** |
| 完全尺度等值 | 11.930 | <0.001 | 0.563 | 0.503 | 21046.854 | 21194.951 | 21093.338 | 0.120[0.112,0.128] | 0.130 | 52.385 | −0.390 | 0.076 |
| 部分尺度等值 | 3.432 | <0.001 | 0.911 | 0.889 | 20536.848 | 20708.086 | 20590.596 | 0.057[0.048,0.066] | 0.051 | 51.992 | −0.035 | 0.012 |
| | | | | | | "I+ W/C" 模型 | | | | | | |
| 完全形态等值 | 1.616 | 0.007 | 0.982 | 0.972 | 20442.823 | 20664.969 | 20512.550 | 0.029[0.015,0.041] | 0.028 | – | – | – |
| **完全单位等值** | **1.551** | **0.010** | **0.983** | **0.975** | **20438.728** | **20642.362** | **20502.644** | **0.027[0.014,0.039]** | **0.031** | 0.868 | **0.001** | **−0.002** |
| 完全尺度等值 | 11.984 | <0.001 | 0.606 | 0.500 | 20986.924 | 21162.790 | 21042.124 | 0.121[0.112,0.129] | 0.138 | 91.970 | −0.377 | 0.094 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 部分尺度等值 | 2.634 | <0.001 | 0.947 | 0.926 | 20490.029 | 20689.035 | 20552.492 | 0.046[0.037,0.057] | 0.041 | 52.470 | −0.036 | 0.019 |

注：G 为一般因子，I 为抑制控制相关因子，W/C 为工作记忆与认知灵活性共同因子。

# 4 讨论

尽管前人不少研究探讨了执行功能的结构和发展，但相关研究绝大多数采用横断设计，并且只关注了执行功能的单个成分，侧重采用单一任务衡量执行功能。基于执行功能的统一性/多样性框架（Miyake & Friedman, 2012），本研究对 756 名童年中期儿童进行间隔一年的追踪，考察了执行功能的结构及稳定性。结果显示，童年中期儿童的执行功能以抑制控制和工作记忆/认知灵活性共同因子的形式（即“I + W/C” 模型）存在；并且 6 项任务被识别为一个社团；统一性和多样性具有一定的稳定性，两者均满足完全单位等值。接下来将对上述研究发现进行详细讨论。

## 4.1 童年中期儿童执行功能的最优结构

个体执行功能的结构会随年龄发生变化（Karr et al., 2018），那么中国文化背景下童年中期儿童执行功能以何种形式存在？本研究在两个时间点上的潜在变量分析发现，单因子（“G”模型）、两因子（“I + W/C”模型、“C + I/W”模型和“I/C + W”模型）以及三因子（“I + W + C”模型）与行为数据拟合良好。进一步模型比较显示，两因子“I + W/C”模型为童年中期的最佳拟合模型，表明该阶段儿童的执行功能以抑制控制和工作记忆与认知灵活性共同因子的形式存在。这意味着抑制控制因素此时已被有效分离出来，但工作记忆和认知灵活性之间具有相同的成分，不可分割（Klauer et al., 2010）。早期前人研究发现，9 岁左右的正常发育儿童中，特定的执行功能彼此之间无法区分；认知灵活性等发展较晚的成分直到 10 ~ 11 岁才可分离出来（Brydges et al., 2014）。本研究证实了前人观察到的这一总体趋势，结果也显示工作记忆和认知灵活性具有共同成分，在该阶段难以分离。可能的原因是：一方面认知灵活性与工作记忆具有更高的相关；另一方面，相较于抑制控制，两者出现相对较晚，因此可能难以将这两者分离出来。此外，本研究采用 1-back 任务和 Digital Span 任务衡量工作记忆，儿童不仅仅是简单地记住当前呈现的信息，还需不断提取前一个或前几个信息，才能获得较好的表现，即涉及记忆的保持和提取的更新过程。同样地，认知灵活性任务也需要更新过程。本研究采用简版 WCST 任务和 TMT 任务衡量认知灵活性，这两项任务同样涉及记忆的保持、提取和更新过程。因此，工作记忆和认知灵活性在本质上都需要在头脑中“提取”上一个信息和“保持”当前信息，这可能是两个成分共享的过程，即工作记忆/认知灵活性共同因子。

## 4.2 衡量童年中期儿童执行功能“统一性和多样性”的指标

关于童年期儿童执行功能的操作性定义，多数研究者通过采用多个任务提取单一成分，将其视为单一结构（e.g., Blair et al., 2011; Clark et al., 2013; Sulik et al., 2015; Vrantsidis et al., 2019）；或采用多个任务提取多个独立成分/潜在变量（e.g., Matte-Gagné et al., 2018; Simanowski & Krajewski, 2019），对执行功能的统一性或多样性进行单独探讨。Teuber（1972）最早在“前额皮层功能的统一性和多样性”一文中提出执行功能的统一性和多样性，并在执行功能的个体差异研究中获得验证。在本研究中，潜在变量分析显示，“I+ W/C”模型为童年中期儿童执行功能的最优模型；然而，网络分析显示，在两个时间点上执行功能 6 项任务指标均被识别为一个社团，凸显出该阶段执行功能结构的统一性。两种分析结果互为补充，强调了统一性和多样性并不是非此即彼，在儿童发展中均发挥重要作用，执行功能的研究需从统一性和多样性两方面同时考察。

起初，关于执行功能的统一性和多样性通过相关性进行评估（Miyake et al., 2000），在相关因子模型中，统一性和多样性反映在因子相关性上。即当因子之间的相关性大于 0，说明存在统一性；而小于 1.0 表明存在多样性。本研究中，衡量执行功能的 6 项任务之间虽然存在显著的相关，但是两两之间的相关性并不一致地较低。评估相同执行功能成分的任务，彼此之间表现出较强的、显著的相关性；评估不同执行功能成分的任务，彼此之间的相关性不那么强，显示出一些收敛和鉴别效度的迹象。此外，不论是单因子"G"模型，还是两因子"I/W + C"模型、"I + W/C"模型和"I/C + W"模型，亦或是三因子"I + W + C"模型，均为行为数据拟合较好；成分相关因子之间的相关性并未达到 1.00，与已有研究发现一致（Feng et al., 2022; Friedman & Miyake, 2017; Miyake et al., 2000, 2001），表明了执行功能具有统一性和多样性。

认知技能的发展具有连续性和层次性，与执行功能直接相关的高级认知技能是在更为简单的认知技能基础上发展起来的（Garon et al., 2008）。而更为简单的认知技能是有效执行各种执行功能任务的必要条件。因此，有必要同时关注更为简单的认知技能（即"统一性"）和与执行功能直接相关的认知技能（即"多样性"）的作用。Friedman 等人（2008）通过构建双因子模型，估计了一个替代相关性的参数——潜在变量，即模型中的执行功能共同潜在变量表明了统一性，特异因子体现了多样性。然而，Friedman 等人未考虑到执行功能结构的年龄差异。相关研究（Brydges et al., 2014; 一项元分析见 Karr et al., 2018）及本研究一致发现，双因子模型在童年期拟合较差，这可能意味着在童年中期个体执行功能还处于发展阶段，并未分化出特异的成分相关因子结构。部分研究者采用因子分析的方法，从多个任务中提取不同的成分，作为统一性和多样性的指标（Friedman et al., 2011; Huizinga et al., 2006; Lehto et al., 2003; Miyake et al., 2000, 2001; Wiebe et al., 2008）。因此，基于潜在变量分析和网络分析的结果，本研究选取"G"模型中的一般因子作为统一性的指标，而最优拟合模型——"I + W/C"模型中的成分相关因子作为多样性的指标。

**4.3 童年中期儿童执行功能统一性和多样性的稳定性**

在探讨执行功能结构的现有研究中，大部分为横断设计，少数关注执行功能发展的研究也只是招募多个年龄群组进行比较（e.g., Davidson et al., 2006; Ferguson et al., 2021; Hartung et al., 2020; Huizinga et al., 2006; Laureys et al., 2022; Lee et al., 2013; Lehto et al., 2003; Xu et al., 2013）。据我们所知，仅有四篇研究采用纵向设计，并进行了纵向等值性以验证执行功能相关结构的稳定性。Usai 等人（2014）对 5 ~ 6 岁儿童（$N = 145$）的执行功能进行间隔一年追踪，发现两因子模型中的"I + W/C"模型在两个时间点上均达到较好拟合，并且纵向等值性检验满足部分尺度等值；Brydges 等人（2014）关注童年中后期执行功能的差异性，对 8 ~ 9 岁儿童（$N = 135$）进行两年追踪，纵向等值性检验仅满足形态等值，表明调查期间因子结构发现变化，T1 时单因子模型拟合较好，T2 时"W + I/C"模型拟合较好。但上述两项研究的样本量较小，对于纵向等值性检验并不友好。Friedman 等人（2016）对 786 名同性双胞胎从青少年后期追踪到成年早期，发现相关因子模型中的三因子模型（即"I + W + C"模型）和双因子模型中的"C + U + S"模型（对应本研究中的"G + W + C"模型）拟合较好，然而仅相关因子模型中的三因子模型满足单位等值；Blair 等人（2014）对 1,292 名 36 个月儿童的执行功能进行追踪调查，发现单因子模型（即"G"模型）满足部分单位等值。上述四项研究考察了不同的年龄阶段；只有一项研究满足部分尺度等值，无法为每个阶

段执行功能结构的稳定性提供强有力的证据；只关注了统一性或多样性的某一方面，未从执行功能的整合视角同时考察统一性和多样性。本研究采用多组验证性因子分析，对相关因子模型中"G"模型和最佳拟合"I + W/C"模型进行纵向等值性检验，发现"I + W/C"模型和"G"模型满足完全单位等值，表明上述结构在测试期间未发生变化，较为稳定，与当前为数不多的两个时间点研究一致（Brydges et al., 2014; Usai et al., 2014），这也提示了执行功能的统一性和多样性在该阶段的稳定性。本研究为当前国内首项从两个时间点上检验执行功能统一性和多样性纵向等值性的研究，不仅揭示了童年中期儿童执行功能的结构，为该结构提供了可靠的证据；也强调了同时考察童年中期儿童执行功能的统一性和多样性的必要性。

值得注意的是，当前研究者对于执行功能的一般因子的含义还存在争论。双因子"G + C + W"模型被众多研究证实是执行功能的最佳模型（e.g., Feng et al., 2022; Friedman et al., 2011, 2016; Glisky et al., 2021; Gustavson et al., 2015; Valian, 2015），但该模型无法分离出抑制控制特异因子。许多执行功能任务需要某种抑制（如，抑制反应、分散、记忆表征等），由此，有部分研究者认为，执行功能一般因子就是抑制控制因子（e.g., Valian, 2015）。然而，这种抑制的概念可能过于宽泛，将概念上和经验上可分离的过程放在一起（Friedman & Miyake, 2004, 2017）。Friedman 和 Miyake（2017）认为，执行功能一般因子反映了个体在保持和管理目标能力上的差异，目标保持和执行是所有执行功能任务的一般要求。个体使用目标来校正正在进行的任务，保持相关信息，该观点目前正获得愈来愈多研究的认可（e.g., Feng et al., 2022）。

不可否认的是，本研究结果对于执行功能结构发展的认识仍具有一定的局限性。一方面，本研究基于执行功能统一性/多样性框架，结合执行功能的三个常见成分选取相关任务进行评估，结果可能受执行功能成分及相关任务的影响。还需进一步了解执行功能结构的组成因子，并确定如何测量这些成分的特异性；另一方面，本研究只考察了童年中期儿童执行功能的结构（小学三至四年级），这在一定程度上限制了对执行功能结构发展趋势的全局描绘。为以更好地了解该阶段执行功能的结构特征，需选取更具代表性的样本，进行更大规模的验证性工作，以进一步扩展当前研究的结果。

## 5 研究结论

中国文化背景下童年中期儿童的执行功能以抑制控制和认知灵活性/工作记忆共同因子（即"I + W/C"模型）的形式存在，并且执行功能的统一性（一般因子）和多样性（成分相关因子）相对稳定。

**参考文献**


Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Blair, C., Protzko, J., & Ursache, A. (2011). Self-regulation and the development of early literacy. In: D. Dickinson, S. Neuman (Eds.), *Handbook of early literacy research* Vol. 3 (pp. 20–35). New York: Guilford Press.

Blair, C., Raver, C. C., Berry, D. J., & The Family Life Project Investigators. (2014). Two approaches to estimating the effect of parenting on the development of executive function in early childhood. *Developmental Psychology*, *50*(2), 554–565. https://doi.org/10.1037/a0033647

Brydges, C. R., Fox, A. M., Reid, C. L., & Anderson, M. (2014). The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis. *Intelligence, 47,* 34–43. https://doi.org/10.1016/j.intell.2014.08.010

Cai, Y., Dong, S., Yuan, S., & Hu, C. (2020). Network analysis and its applications in psychology. *Advances in Psychological Science*, *28*(01), 178–195.

[蔡玉清,董书阳,袁帅，胡传鹏. (2020). 变量间的网络分析模型及其应用. *心理科学进展*, *28*(01), 178–195.]

Carlson, S. M., White, R. E., & Davis-Unger, A. C. (2014). Evidence for a relation between executive function and pretense representation in preschool children. *Cognitive Development, 29,* 1–16. https://doi.org/10.1016/j.cogdev.2013.09.001

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Clark, C. A. C., Sheffield, T. D., Wiebe, S. A., & Espy, K. A. (2013). Longitudinal associations between executive control and developing mathematical competence in preschool boys and girls. *Child Development*, *84*(2), 662–677. https://doi.org/10.1111/j.1467-8624.2012.01854.x

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the Cantab. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology*, *25*(2), 242–254. https://doi.org/10.1076/jcen.25.2.242.13639

Downes, M., Bathelt, J., & De Haan, M. (2017). Event-related potential measures of executive functioning from preschool to adolescence. *Developmental Medicine & Child Neurology*, *59*(6), 581–590. https://doi.org/10.1111/dmcn.13395

Epskamp, S. (2015). *bootnet: Bootstrap methods for various network estimation routines. R-Package*.

Espy, K. A., Kaufmann, P. M., McDiarmid, M. D., & Glisky, M. L. (1999). Executive functioning in preschool children: Performance on A-not-B and other delayed response format tasks. *Brain and Cognition, 41*(2), 178–199. https://doi.org/10.1006/brcg.1999.1117

Feng, J., Zhang, L., Chen, C., Sheng, J., Ye, Z., Feng, K., Liu, J., Cai, Y., Zhu, B., Yu, Z., Chen, C., Dong, Q., & Xue, G. (2022). A cognitive neurogenetic approach to uncovering the structure of executive functions. *Nature Communications, 13*, Article 4588. https://doi.org/10.1038/s41467-022-32383-0

Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports*, *11*, Article 1382. https://doi.org/10.1038/s41598-020-80866-1

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Friedman, N. P., Miyake, A., Altamirano, L. J., Corley, R. P., Young, S. E., Rhea, S. A., & Hewitt, J. K. (2016). Stability and change in executive function abilities from late adolescence to early adulthood: A longitudinal twin study. *Developmental Psychology*, *52*(2), 326–340. https://doi.org/10.1037/dev0000075

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology*, *47*(5), 1410–1430. https://doi.org/10.1037/a0023750

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General, 137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Glisky, E. L., Alexander, G. E., Hou, M., Kawa, K., Woolverton, C. B., Zigman, E. K., Nguyen, L. A., Haws, K., Figueredo, A. J., & Ryan, L. (2021). Differences between young and older adults in unity and diversity of executive functions. *Aging, Neuropsychology, and Cognition*, *28*(6), 829–854. https://doi.org/10.1080/13825585.2020.1830936

Goldstein, S., & Naglieri, J. A. (2014). *Handbook of executive functioning.* New York: Springer.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Gustavson, D. E., Miyake, A., Hewitt, J. K., & Friedman, N. P. (2015). Understanding the cognitive and genetic underpinnings of procrastination: Evidence for shared genetic influences with goal management and executive function abilities. *Journal of Experimental Psychology: General*, *144*(6), 1063–1079. https://doi.org/10.1037/xge0000110

Hartung, J., Engelhardt, L. E., Thibodeaux, M. L., Harden, K. P., & Tucker-Drob, E. M. (2020). Developmental transformations in the structure of executive functions. *Journal of Experimental Child Psychology*, *189*, Article 104681. https://doi.org/10.1016/j.jecp.2019.104681

Hu, L.−T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hughes, C., & Devine, R. T. (2019). For better or for worse? Positive and negative parental influences on young children's executive function, *Child Development*, *90*(2), 593–609. https://doi.org/10.1111/cdev.12915

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017–2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

IBM Corp. *Released 2020. IBM SPSS Statistics for Windows, Version 20.0*. Armonk, NY: IBM Corp.

Isquith, P. K., Roth, R. M., & Gioia, G. A. (2013). Contribution of rating scales to the assessment of executive functions. *Applied Neuropsychology: Child*, *2*(2), 125–132. https://doi.org/10.1080/21622965.2013.748389

JASP Team. (2022). *JASP (Version 0.16.4) [Computer software]*. Retrieved from https://jasp−stats.org/.

Jung, E., & Yoon, M. (2016). Comparisons of three empirical methods for partial factorial invariance: Forward, backward, and factor−ratio tests. *Structural Equation Modeling*, *23*(4), 567–584. https://doi.org/10.1080/10705511.2015.1138092

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

Klauer, K. C., Schmitz, F., Teige-Mocigemba, S., & Voss, A. (2010). Understanding the role of executive control in the implicit association test: Why flexible people have small IAT effects. *Quarterly Journal of Experimental Psychology, 63*(3), 595–619. https://doi.org/10.1080/17470210903076826

Lange, F., Seer, C., Müller-Vahl, K., & Kopp, B. (2017). Cognitive flexibility and its electrophysiological correlates in Gilles de la Tourette syndrome. *Developmental Cognitive Neuroscience*, *27*, 78–90. https://doi.org/10.1016/j.dcn.2017.08.008

Laureys, F., De Waelle, S., Barendse, M. T., Lenoir, M., & Deconinck, F. J. (2022). The factor structure of executive function in childhood and adolescence. *Intelligence*, *90*, Article 101600. https://doi.org/10.1016/j.intell.2021.101600

Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development*, *84*(6), 1933–1953. https://doi.org/10.1111/cdev.12096

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*(1), 59–80. https://doi.org/10.1348/026151003321164627

Maasalo, K., Lindblom, J., Kiviruusu, O., Santalahti, P., & Aronen, E. T. (2021). Longitudinal associations between inhibitory control and externalizing and internalizing symptoms in school-aged children. *Development and Psychopathology*, *33*(3), 843–855. https://doi.org/10.1017/s0954579420000176

Matte-Gagné, C., Bernier, A., Sirois, M.-S., Lalonde, G., & Hertz, S. (2018). Attachment security and developmental patterns of growth in executive functioning during early elementary school. *Child Development*, *89*(3), e167–e182. https://doi.org/10.1111/cdev.12807

McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society*, *16*(3), 495–505. https://doi.org/10.1017/S1355617710000093

McCoy, D. C. (2019). Measuring young children's executive function and self-regulation in classrooms and other real-world settings. *Clinical Child and Family Psychology Review*, *22*(1), 63–74. https://doi.org/10.1007/s10567-019-00285-1

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General, 130*(4), 621–640. https://doi.org/10.1037/0096-3445.130.4.621

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide (8th ed.)*. Los Angeles, CA: Author.

Pelegrina, S., Molina, R., Rodríguez-Martínez, E. I., Linares, R., & Gómez, C. M. (2020). Age-related changes in selection, recognition, updating and maintenance information in WM. An ERP study in children and adolescents. *Biological Psychology*, *157*, Article 107977. https://doi.org/10.1016/j.biopsycho.2020.107977

Qi, Y., Yang, G., Fu, D., Li, Z., & Liu, X. (2021). Developmental neuroscience of cognitive control: The future path and layout. *Science Sinica Vitae*, *51*(6), 634–646.

[齐玥, 杨国春, 付迪, 李政汉, 刘勋. (2021). 认知控制发展神经科学：未来路径与布局. *中国科学生命科学*, *51*(6), 634–646.]

Qiu, Z., & Lin, B. (2019). Principle and application of the structural equation model (second edition). Beijing: Chinese light industry press.

[邱皓政, 林碧芳. (2019). 结构方程模型的原理与应用（第二版）. 北京: 中国轻工业出版社.]

R Core Team (2020). *A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Raven, J. C., Court, J. H., & Raven J. (1983). *Manual for Raven's Progressive Matrices and Vocabulary Scales* (Section 4: Advanced Progressive Matrices). London: H. K. Lewis.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507–514. https://doi.org/10.1007/BF02296192

Seer, C., Lange, F., Georgiev, D., Jahanshahi, M., & Kopp, B. (2016). Event-related potentials and cognition in Parkinson's disease: An integrative review. *Neuroscience and Biobehavioral Reviews, 71,* 691–714. https://doi.org/10.1016/j.neubiorev.2016.08.003

Shi, R., Sharpe, L., & Abbott, M. (2019). A meta-analysis of the relationship between anxiety and attentional control. *Clinical Psychology Review, 72,* Article 101754. https://doi.org/10.1016/j.cpr.2019.101754

Simanowski, S., & Krajewski, K. (2019). Specific preschool executive functions predict unique aspects of mathematics development: A 3-year longitudinal study. *Child Development*, *90*(2), 544–561. https://doi.org/10.1111/cdev.12909

Song, Z., Guo, L., & Zheng, T. (2022). Comparison of missing data handling methods in cognitive diagnosis: Zero replacement, multiple imputation and maximum likelihood estimation. *Acta Psychologica Sinica*, *54*(4), 426−440.

[宋枝璘, 郭磊, 郑天鹏. (2022). 认知诊断缺失数据处理方法的比较：零替换、多重插补与极大似然估计法. *心理学报*, *54*(4), 426−440. ]

Strauss, E., Sherman, E. M. X., & Spreen, O. (2006). *A compendium of neuropsychological tests* (3rd edn). New York: Oxford University Press.

Sulik, M. J., Blair, C., Mills-Koonce, R., Berry, D., Greenberg, M., & The Family Life Project Investigators. (2015). Early parenting and the development of externalizing behavior problems: Longitudinal mediation through children's executive function. *Child Development*, *86*(5), 1588–1603. https://doi.org/10.1111/cdev.12386

Takacs, Z. K., & Kassai, R. (2019). The efficacy of different interventions to foster children's executive function skills: A series of meta-analyses. *Psychological Bulletin*, *145*(7), 653–697. https://doi.org/10.1037/bul0000195

Teuber, H. L. (1972). Unity and diversity of frontal lobe functions. *Acta Neurobiologiae Experimentalis*, *32*(2), 615–656.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*(2), 131–143. https://doi.org/10.1111/jcpp.12001

Usai, M. C., Viterbori, P., Traverso, L., & De Franchis, V. (2014). Latent structure of executive function in five- and six-year-old children: A longitudinal study. *European Journal of Developmental Psychology, 11*(4), 447–462. https://doi.org/10.1080/17405629.2013.840578

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18,* 3–24.

Vrantsidis, D., Clark, C. A. C., Chevalier, N., Espy, K. A., & Wiebe, S. (2019) Socioeconomic status and executive function in early childhood: Exploring proximal mechanisms, *Developmental Science*, *23*, Article 12917. https://doi.org/10.1111/desc.12917

Weschler. (2002). *Wechsler Preschool and Primary Scale of Intelligence - Third Edition*. Harcourt Assessment.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*(2), 575–587. https://doi.org/10.1037/0012-1649.44.2.575

Wild, K. V., & Musser, E. D. (2014). The Cambridge neuropsychological test automated battery in the assessment of executive functioning. In: S. Goldstein and J. A. Naglieri (eds.), *Handbook of executive function* (pp.171–190). New York: Springer.

Xu, F., Han, Y., Sabbagh, M. A., Wang, T., Ren, X., & Li, C. (2013). Developmental differences in the structure of executive function in middle childhood and adolescence. *PLoS ONE*, *8*(10), Article 77770. https://doi.org/10.1371/journal.pone.0077770

Yaple, Z., & Arsalidou, M. (2018). N-back working memory task: Meta-analysis of normative fMRI studies with children. *Child Development*, *89*(6), 2010–2022. https://doi.org/10.1111/cdev.13080

Younger, J. W., O'Laughlin, K. D., Anguera, J. A., Bunge, S. A., Ferrer, E. E., Hoeft, F., McCandliss, B. D., Mishra, J., Rosenberg-Lee, M., Gazzaley, A., & Uncapher, M. R. (2023). Better together: Novel methods for measuring and modeling development of executive function diversity while accounting for unity. *Frontiers in Human Neuroscience*, 17, Article 1195013. https://doi.org/10.3389/fnhum.2023.1195013

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*(4), 354–360. https://doi.org/10.1111/j.1750-8606.2012.00246.x

Zhang, H., & Wang, X. (1989). Standardization research on Raven's standard processive matrices in China. *Acta Psychologica Sinica*, *21*(2), 113–121.

[张厚粲, 王晓平. (1989). 瑞文标准推理测验在我国的修订. *心理学报*, *21*(2), 113–121.]

（通讯作者邮箱：司继伟 Email: sijiwei@126.com）

作者贡献声明 ：

黄碧娟，司继伟：提出研究思路，设计研究方案；

黄碧娟，朱红祥，刘畅，封洪敏，雒瑞帆：进行实验；

黄碧娟，李红霞：采集、清洗和分析数据；

黄碧娟：论文起草；

黄碧娟，司继伟：论文最终版本修订。